

## DOCUMENT RESUME

ED 126 490

CS 002 855

AUTHOR Piontkowski, Dorothy  
TITLE Clinical Assessment of Beginning Reading Skills.  
PUB DATE. 76  
NOTE 14p.; Paper presented at the Annual Meeting of the  
International Reading Association (21st, Anaheim,  
California, May 1976)

EDRS PRICE MF-\$0.83 HC-\$1.67 Plus Postage.  
DESCRIPTORS Basic Skills; \*Beginning Reading; Elementary  
Education; Grade 1; Reading Development; \*Reading  
Processes; Reading Research; \*Reading Skills;  
\*Reading Tests; \*Test Construction

## ABSTRACT

In order to understand better the elemental processes involved in beginning reading (specifically, how cognitive skills are acquired and whether a relationship exists among basic skills), Robert Calfee and colleagues conducted a first-grade reading diary study. Forty target children in 1973-1974 and 50 children in 1974-1975 were observed as they were being taught to read and were given systematic and continual tests designed according to the development of their reading skills. This paper describes the construction of the network of tests which were used for the study and which were designed to identify the specific skills directly related to reading acquisition and to control for those general abilities which influence a student's performance on virtually any task. Although data from the study have not been completely analyzed, it is apparent both that the elemental processes being measured did provide the basis of beginning reading achievement and that this method of assessment offers teachers precise diagnostic information about the instructional needs of beginning students. (JM)

\*\*\*\*\*  
\* Documents acquired by ERIC include many informal unpublished \*  
\* materials not available from other sources. ERIC makes every effort \*  
\* to obtain the best copy available. Nevertheless, items of marginal \*  
\* reproducibility are often encountered and this affects the quality \*  
\* of the microfiche and hardcopy reproductions ERIC makes available \*  
\* via the ERIC Document Reproduction Service (EDRS). EDRS is not \*  
\* responsible for the quality of the original document. Reproductions \*  
\* supplied by EDRS are the best that can be made from the original. \*  
\*\*\*\*\*

## Clinical Assessment of Beginning Reading Skills

Dorothy Piontkowski  
Stanford University

ED126490

A case history study has the potential of providing a wealth of longitudinal information not available in other designs. The First Grade Reading Diary Study is a unique example of a case history approach aimed toward providing description of how children acquire beginning reading skills. This study was carried out under the auspices of the Carnegie Corporation by Dr. Robert Calfee and his colleagues at Stanford University. It aimed toward a better understanding of the elemental processes in beginning reading. How are these cognitive skills acquired? Is there a relationship among basic reading skills?

Forty target children in 1973-74 and 50 children in 1974-75 from first grade classrooms in the San Jose Unified School District were observed while they were being taught to read, and were given systematic and continual tests designed according to the development of their reading skills. The classrooms were selected from low to moderate income neighborhood schools using teachers with a variety of classroom philosophies, teaching styles, and curriculum choices. Variations in target classrooms, and in teaching styles and method, were included to isolate the effects of instruction used to promote acquisition of decoding skills. Students were selected from among those who we predicted would learn to read, but were not already doing so upon entering first grade. The children represented a range of ethnic backgrounds.

Two primary data collecting systems were used: (1) a comprehensive observational system designed to capture the historical events within the classroom and (2) a network of tests constructed to yield precise measures

of specific, componential skills acquired along the path toward reading. Observational systems generally employed have tended to be narrowly confined to limited or sporadic classroom activities, to be insensitive to behavioral frequencies, to be tied to a priori categories, or to be built upon observer-defined scales of subjective indices. For our case history approach, a comprehensive system of objective observation, RAMOS (Reading and Mathematics Observation System [Calfee & Hoover, 1974]), was utilized to provide a computer-compatible summary of classroom activity and structure. Observations took place weekly during reading instruction and focused on many features in each of the classrooms. The flexibility of RAMOS categories permitted analyses of the activities of the target children, the functioning of the teacher, the usage of space, the content of and time allotted to instruction, the teacher's feedback procedures, and a number of other coded behaviors and situations. These features of RAMOS enabled us to accumulate an objective, historical description of what was going on in each first grade classroom throughout the school year.

Concurrent with the accumulation of descriptive observations, the Reading Diary Study determined precisely what each student learned from week to week, and to what extent this knowledge had generalized. This challenge required tests that would pinpoint separable components along the continuum from nonreading to reading.

Tests currently used for assessment of reading achievement possess great predictive reliability, but are not sensitive to the elemental processes underlying mastery of larger skills. Most existing tests are constructed so that in order to do well a student must bring to bear a complex variety of skills and knowledge. For our purposes, we needed specific kinds of information quite independent of norm-referenced performance that

mask precise individual differences in basic competencies. Our goal was to identify the specific skills directly related to reading acquisition and to control for those general abilities that influence a student's performance on virtually any tasks, confounding his competence on elemental tasks.

General abilities include knowledge of how a task is to be performed, a set toward speed rather than accuracy, and availability of coding systems for processing incoming information and reducing short-term memory load.

We developed an assessment system to obtain what Glaser (1970) has called "the outcomes of learning"--the specific performance-skills of the target students on precisely defined measures. The underlying rationale is in line with Gagné's (1970) description of distortion-free measurement in which performance is not influenced by any variable other than the one explicitly defined. Such tests are "clean" (Calfee, 1976)--the task requirements are simplified, the relevant cues are salient, and the substance of the task is designed to match the student's past learning experiences as nearly as possible.

Before examining the hierarchical structure across subtests, we will focus on the explicit controls for contamination within each test-cell or subtest. In order to direct the child's attention to the relevant cues in each subtest, two sample items were provided with corrective feedback immediately prior to the subtest. These sample items helped clarify directions by example. To control for a child's set toward speed rather than accuracy, all testing was done individually with the pace determined by the tester. More than one correct item was frequently included among each multiple-choice array in order to stress inclusive search-strategies. The boredom and fatigue brought on by lengthy, repetitive subtests were avoided by succinct test-cells of only four items that quickly indicated mastery or ignorance of a specific skill.

At the same time, we developed a system to avoid presenting subtests beyond the competence of the individual student, or taking the more skilled student through a lengthy series of subtests that measured skills he had already mastered. A dynamic or clinical format was created which allowed the tester to plot each child's course through the subtests on which he was successful while minimizing his exposure to subtests that were too easy or too difficult for him. Wittrock (1970) has referred to this as a "multistage decision process." The tester follows a prescribed branching rule making decisions sequentially which affect subsequent decisions. The length and sequence of testing varies according to each child's performance, and continues until the branching rule cannot be satisfied, that is, until the child has mastered the most difficult subtests, and his level of competence has been completely defined by the intersection of subtests mastered and subtests failed. In this way, untested skills are assumed to have been mastered or not yet learned according to their position in the hierarchy and the student's responses.

The initial phase of test construction was based on a theoretical sketch of the logical classification system in which the concepts or beginning reading are generally taught. Thus, the sequence and structure of the assessment scheme reflect the range of skills directly related to instruction. Analyses of methods and materials allowed for predicting the steps a child takes as he progresses from nonreader to beginning reader. Within this series, two subgoals are apparent: (1) the child learns an alphabetic foundation of the written system and letter-sound correspondences in order to analyze and synthesize the letter combinations that lead to reading, and (2) the child learns to read sequential strings of words to finally arrive at meaning and comprehension. Within each of these subgoals exist the elemental processes and skills that must be mastered in learning to read.

The first step toward definition of the elemental processes was an analysis of the essential attributes of the content of beginning reading leading to development of a general classification scheme for the assessment network. The primary substantive variable appeared to be orthographic composition. We hypothesized that word readability is a serial progression of difficulty starting with consonant-vowel-consonant words, and followed by consonant blends and digraphs, vowel contrasts, and polysyllabic words. This separation of words by spelling pattern defined the primary categoric boundary, and specified which word-items could be included in each test-scheme within the network while explicitly excluding all other word-items. Figure 1 provides a diagram of the primary dimension of the testing network.

A multi-dimensional design of subtests was generated according to the inherent dimensions of variability within the content and form of each word-category. That is, each category was analyzed into dimensions according to the structural boundaries imposed by the orthography of the words, and by the tasks which represent the specific skills requisite to mastery of that category. In general, we included and extended the guidelines proposed by Baker (1974). One dimension consists of all the eligible content specifications: all areas of separable content required to assure master of that category of letter-sound correspondence. Another dimension is concerned with the form in which the learner must display his knowledge, all ways the child might demonstrate those behaviors indicating mastery. A third dimension measures the child's ability to generalize his specific knowledge to materials differing in degree of temporal availability and familiarity. Figure 2 presents the three dimensional model of the consonant-vowel-consonant test scheme.



Each of these content and response dimensions were then factored into the significant elements of that dimension. Underlying the ability to read the whole word, the child needs to learn how to deal with specific components of that word--the initial consonant, final consonant, medial vowel. Within the dimension called "letter-sound environment" we generated measures of each of these elemental processes. We factored "task" variables in order to assess differentially the child's ability to recognize the correct response from an array of letters or words, to produce verbally the correct response, and to write the appropriate letter or word. In order to measure transfer of knowledge about consonant-vowel-consonant words, we factored "materials" into three elements: pictured and verbalized stimuli, verbalized labels, and nonsense words. The three-dimensional framework of the consonant-vowel-consonant category is presented in a nested scheme in Figure 3.

Each subtest is defined by the crossing of three specified elements within the three dimensions of form and content. Humphreys (1962) and Guttman (1969) have defined this method of test construction as subtest definition by explicit specification of the combinations of elements. Subtests are generated across the dimensions by specific limitations on the items for each subtest according to the restrictions imposed by the structural definitions. New cells are created by holding two dimensional elements constant and varying only the third. Each item within each subtest must fit the precise factorial definition. By moving a child through the subtests, systematic sources of difficulty are added and the child's performance indicates the precise situational alteration that separates his mastery from ignorance. In this way, each testing session determines the student's knowledge of beginning reading defined by the factorial dimensions of that test. By starting with subtests

yielding low error frequency and systematically introducing sources of difficulty unidimensionally, the child's performance across subtests provides a diagnostic write-up of his elemental skills.

Thus, the test structure per se is an integral part of the entire instrument; the structure of the test defines precisely what it is the child can and cannot do. The multidimensional network of subtests mirror the elemental subskills that are required for mastery of the greater skill. Each subtest measures a factorially unique process so that when the nested subskills are combined again, they become that skill. When taken separately, each subtest provides diagnosis within definition of the dimensional elements; when grouped, the subtests provide a measure of the child's knowledge across elements and dimensions. By examining the child's performance, the tester is able to arrive at an unbiased assessment of the child's specific abilities on an explicitly defined hierarchy of reading skills.

In addition to the multi-dimensional test-schemes for word-categories, entry and probing systems surround each scheme. Figure 4 provides a diagrammatic plan of the more extensive network. Entry tests were constructed according to the word-item limitations imposed by the word-category, and in addition, contained irregular functional words to assess the child's ability to recognize and read certain high frequency words not included in any of the categorical schemes. The entry tests did not measure analytic abilities, but were limited to two performance tasks: identification of a verbalized stimuli from a printed array, and ability to read a list of words. The entry tests were vehicles for making branching decisions about which test-scheme to present to the child, and at what level of difficulty to enter.

-----  
Figure 4  
about here  
-----



A comprehensive system of probing tests was developed surrounding the consonant-vowel-consonant test scheme in order to hone in on specific sources of difficulty. These included tests of letter-sound correspondences, letter naming, letter writing, and letter discrimination.

As indicated across the top of Figure 4, we also tested the target students' ability to read prose derived from words of similar orthographic patterns to "fun" picture books. In two of the prose reading tests, comprehension questions were asked at the conclusion of the oral reading. Although the data have not been completely analyzed as of this time, it is apparent that those children unable to perform the analytic and synthetic tasks within the test-schemes were also not able to read the connected prose. Apparently the elemental processes we were seeking to measure did provide the basis of beginning reading achievement.

The low correlation between scores on the subtests and the bimodality of performances on subtests indicate that we may have approached the elemental level of processes within our test-schemes. Doing well on one subtest did not predict success as additional sources of difficulty were added. We were able to arrive at and identify the precise source of difficulty beyond which the child could not succeed. This manner of testing identifies specific levels of mastery or ignorance for individual students, and has the power to identify areas in which a beginning reading program needs modification in order for some children to succeed. Likewise, those students already mastering elemental processes may be identified and moved on to more appropriate learning situations. This method of assessment offers promise of factoring out the significant dimensions of beginning reading behaviors and of offering the teacher precise diagnostic information about the instructional needs of beginning readers.

# References

- Calfee, R. C. Assessment of independent reading skills: Basic research and practical applications.. In A. S. Reber & D. Scarborough (Eds.), Toward a psychology of reading: The proceedings of CUNY Conference, Hillsdale, N. J.: Lawrence Erlbaum Associates, Inc., 1976.
- Calfee, R. C., & Hoover, K. Reading and mathematics observation system. Unpublished paper, Stanford University, 1974.
- Baker, E. L. Beyond objectives: Domain-referenced tests for evaluation and instructional improvement. In W. Hively (Ed.), Domain-referenced testing. Englewood Cliffs, N. J.: Educational Technology Publications, 1974.
- Gagné, R. Instructional variables and learning outcomes. In M. C. Wittrock & D. E. Wiley (Eds.), The evaluation of instruction: Issues and problems. New York: Holt, Rinehart and Winston, Inc., 1970.
- Glaser, R. Evaluation of instruction and changing educational models. In M. C. Wittrock & D. E. Wiley (Eds.), The evaluation of instruction: Issues and problems. New York: Holt, Rinehart and Winston, Inc., 1970.
- Guttman, L. Integration of test design and analysis. Invitational conference on testing. Educational Testing Service, 1969.
- Humphreys, L. The organization of human abilities. American Psychologist, 1962, 17(7), pp. 475-483.
- Wittrock, M. C. The evaluation of instruction: Cause-and effect relations in naturalistic data. In M. C. Wittrock & D. E. Wiley (Eds.), The evaluation of instruction: Issues and problems. New York: Holt, Rinehart and Winston, Inc., 1970.

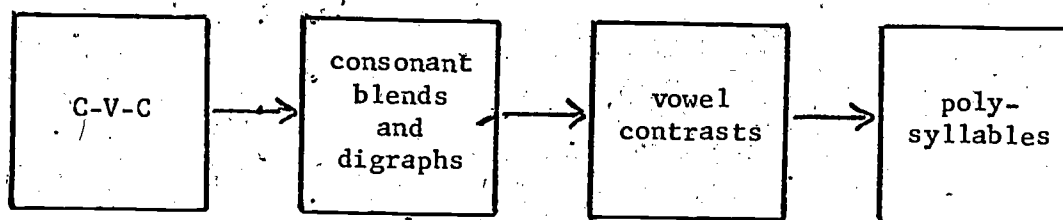


Figure 1. The primary classification system for the assessment network based on orthographic composition of words.

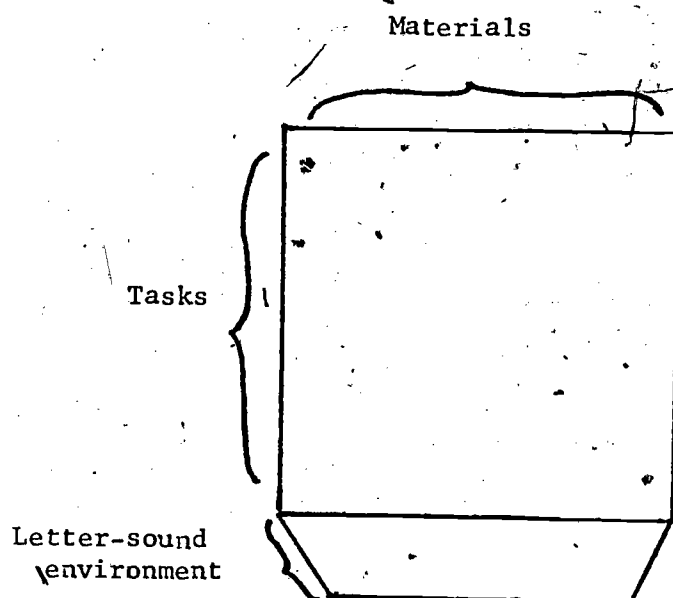


Figure 2. Three-dimensional model of the consonant-vowel-consonant test-scheme developed for the Reading Diary Study.

	Identification			Verbal Production			Written Production		
	picture	verbal	non-word	picture	verbal	non-word	picture	verbal	non-word
initial cons.									
final cons.									
medial vowel									
word									

Figure 3. The three dimensional framework of tasks, materials, and letter-sound environments which created the 36 subtests of the consonant-vowel-consonant test-scheme

# Oral reading tests:

Final vowel-consonant patterns

Four levels of "fun" reading

Basal readers: past, present, future lessons

Paragraphs constructed from words in the test-scheme

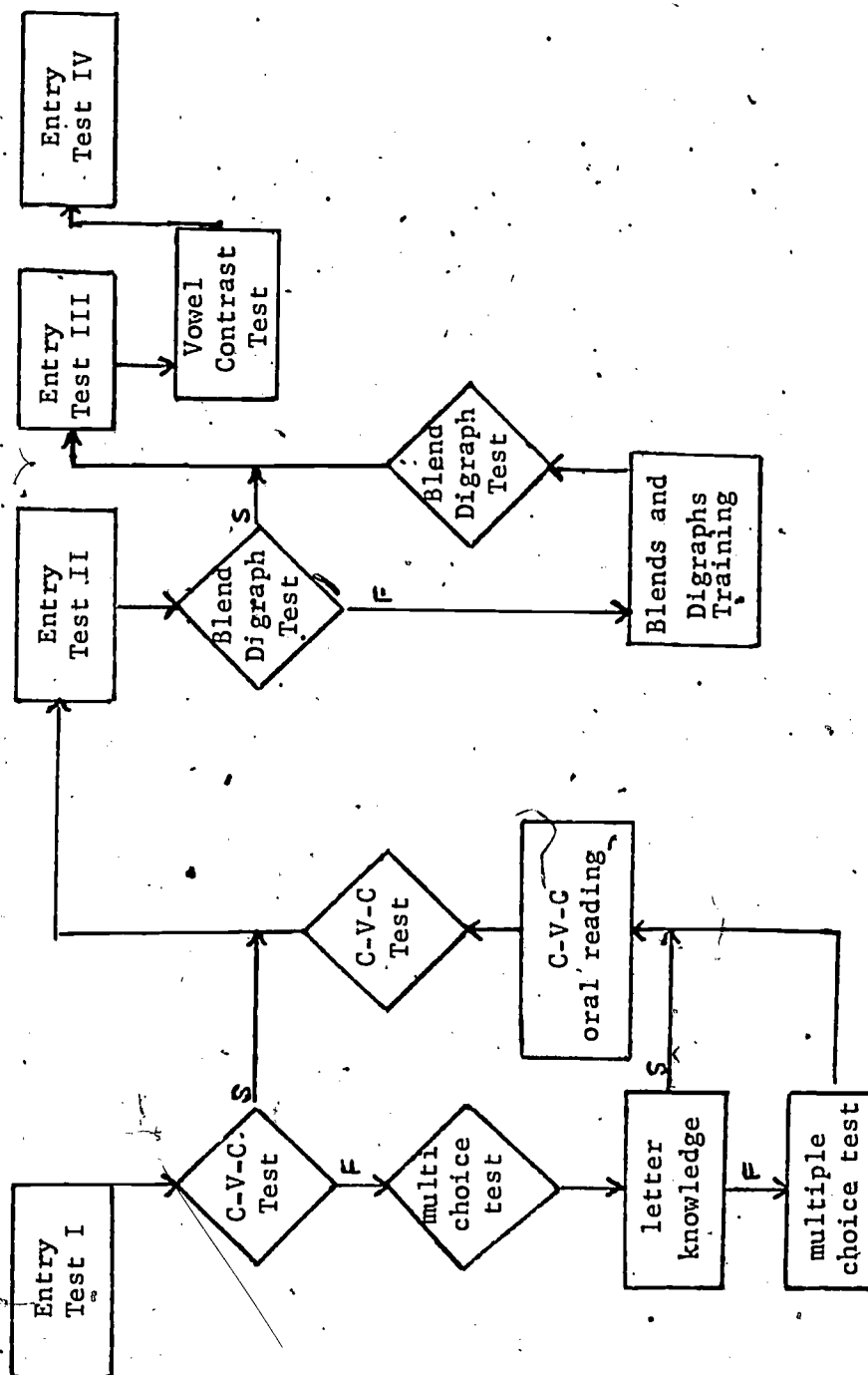


Figure 4. A diagram of the assessment network developed for the Reading Diary Study and the list of oral reading tests used for evaluating the student's skills in beginning reading.